(12) # EUROPEAN PATENT APPLICATION

(21) Application number: 84401120.5

(22) Date of filing: 30.05.84

(51) Int. Cl.³: **H 04 N 7/08**

(71) Applicant: TEXAS INSTRUMENTS FRANCE Société dite:

F-06270 Villeneuve Loubet(FR)

(84) Designated Contracting States:
FR

(71) Applicant: TEXAS INSTRUMENTS INCORPORATED
13500 North Central Expressway
Dallas Texas 75265(US)

(84) Designated Contracting States:
DE GB IT NL

(72) Inventor: Benbassat, Gérard V.
1185 Route de la Colle
F-06570 St Paul(FR)

(72) Inventor: Serain, Daniel
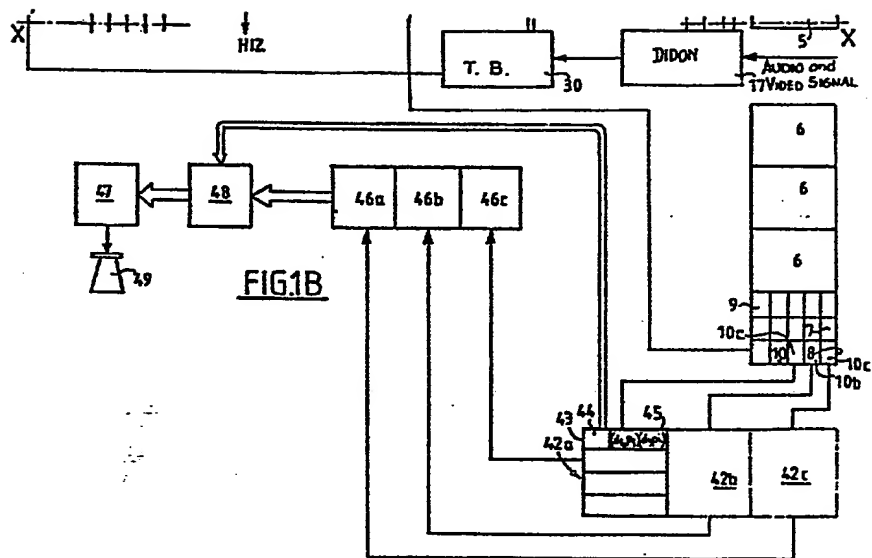56 Avenue Parc Mosca
F-06000 Nice(FR)

(74) Representative: Moncheny, Michel et al,
c/o Cabinet Lavoix 2 Place d'Estienne d'Orves
F-75441 Paris Cedex 09(FR)

(54) A general technique to add multilingual speech to videotext systems at a low data rate.

(57) A device for coordinating the audition of spoken messages with the visualization of messages, formed of pages, on a display apparatus, characterized in that it includes means for memorizing the codes corresponding to a written text to be visualized, means (6) for memorizing the codes corresponding to a spoken message to be reproduced, means (7) for associating the text and the speech corresponding to the codes stored in said memorizing means (6), means (42a to 42c) for memorizing the intonation elements (di, pi) of the sound segments forming the received message, means (46a to 46c) for memorizing the elements constituting the language in which the message is received, means (48) for combining said intonation elements with said elements constituting the language, and means (47) for synthesizing the message to be reproduced under control of said combining means.

EP 0 128 093 A1

./...

X ———|—|—|—|—|——   ↓   [T. B.   ] ←—— [DIDON   ]   ———|—|—|—|—|——— X
                         HIZ   30   AUDIO and   5'
                                                                       17 VIDEO SIGNAL

[47] ← [48] ← [46a | 46b | 46c]

49

FIG.1B

6

6

.

6

9
10c
10   8   10c
                    10b

43 44   45
42a   [42b] [42c]

A general technique to add multilingual speech 0128093
to videotext systems at a low data rate.-

This invention relates to the coordination
of audible signals and images appearing on a display
device and, more specifically, to the audition of
spoken communications corresponding to messages in the
5   form of pages displayed on the screen of a cathode ray
tube.

A process and apparatus for the display of
messages on a sweep frame display device, such as the
screen of a cathode ray tube, are described in French
10   Patent Applications Nos 80 26 392, 80 26 393, 80 26 394,
80 26 395 of the instant applicant.

These patent applications describe a technique
which provides for edition, transmission, reception,
and visualization of messages containing typographic
15   information which can be intermixed with graphical data.

These techniques are particularly applicable
to a teletext system such as the Antiope system.

The invention relates to a speaking or
audible teletext system which adjoins to the graphical
20   and/or typographical data, visible on the screen of a
cathode tube, audible data corresponding thereto such
as, for example, commentaries on a written text, or a
graphical representation, or another.

The invention has, therefore, as an object
25   a process for the coordination of the audition of spoken
messages with the visualization of messages comprising
pages on a display device characterized in that it
includes, receiving the image signals and the word
signals encoded in similar codes, the codes correspon-
30   ding to the written text being recognizable by a first
series of numbers of rows of the page assigned to these
written text codes with the codes corresponding to the
spoken message being recognizable by means of a second
series of numbers of rows of the page assigned to these
35   spoken message codes, effecting the discrimination of

the codes of the written text and the codes of the
spoken message, storing the codes of the written text
in a first group of memories, storing the codes of the
spoken text in a second group of memories, causing the

5   association of the codes of the written text with
the codes of the corresponding spoken message by means
of an association table, reconstituting the spoken
message from the codes contained in the second group
of memories, and causing its audible emission, follo-

10   wing a predetermined sequence, with the display of the
corresponding written text on the screen of a cathode
ray tube.

The invention also has as its object an
apparatus for the coordination of the audition of spoken

15   messages with the visualization of messages composed
of pages on a display device characterized in that it
includes means for memorizing the codes corresponding
to a written text to be displayed, means for memorizing
the codes corresponding to the spoken message to be

20   reproduced, means for associating the text and the
word corresponding to the codes stored in said memory
means,means for memorizing the intonation elements of the
audible segments forming the received message, means
for memorizing the elements constituting the language

25   in which the message is received, means for combining
said intonation elements with said constituting elements,
and means for synthesizing the message to be reproduced
under control of said combining means.

The invention will be better understood with

30   reference to the following description of a particular
embodiment, which is set forth only as an example of
the invention and with reference to the drawings.

Fig. 1a and 1b is a general schematic of a
receiving station, for the display of teletext pages,

- 3 -

of the teletext system "antiope", and for the audition
of sound data corresponding to the teletext pages.

Fig. 2 shows the organization of a page of
text to be reproduced in written and spoken form by
5 means of the receiver of Fig. 1.

Before proceeding with the description of
the invention itself, there will be presented some
generalities as to the structure of the antiope video
text system, to which the invention is applied.

10 This system provides for the transmission
of digital information over non-utilized television
lines.

Upon reception, the binary bits are grouped
into bytes.

15 The antiope system effects the transmission
of characters in the ASCII codes by means of herzian
waves. The transmitted data are structured in the form
of magazines or groups of pages.

A magazine includes generally a front page,
20 a summary page, and information pages, these pages being
numbered from zero to 999.

Each page has a heading setting forth its
characteristics and a group of 25 rows numbered 0 to 24.

These 25 rows of codes correspond to the 25
25 rows of a television screen, each row containing 40
columns.

There are thus 1,000 character positions on
the screen. To each of these positions there corresponds
a word of sixteen binary bits contained in a display
30 memory which identify the character to be displayed as
well as its characteristics (size, color, etc.).

The ASCII code of the displayed character is
utilized as an address by the display processor for
reading into the alphabet memory of the assembly, the
35 description of the form of the character.

It will be noted that this alphabet memory
can be a RAM. In this case, the description of the
characters to be displayed can itself be transmitted
by the television channel in the form of a non-displayable
5   special page. Such a technique allows the changing of
the alphabet in a dynamic manner at the request of
the user.

     This technique is known as "dynamically
redefinable character set" DRCS.

10   Other videotexted systems are of the
"directory" terminal type. These systems utilize the
same data coding system as the antiope system.

     They differ from this system as to two
principal points, these being the interactivity and
15   the transmission support.

     The directory terminal system is interactive.
The user can select the transmitted pages. The user
requests the emitting source to send the pages which
are of interest. The transmission support is a standard
20   telephone line. A modem is necessary for transmitting
the digital data.

     The receptor station seen in Fig. 1 which will
now be described can provide spoken and written
responses upon the user's request.

25   There is included a group of memories or
composite memory 1, with various sections being
assigned to data of different nature.

     There is included a buffer memory 2 for
temporarily storing the entry data with the separated
30   video signals and audio signals applied to its
terminal (arrow 3).

     Such a buffer memory is described in detail
in the above mentioned patent application N° 80 26 393
entitled "Process and Apparatus for the Exchange of
35   Information Between Information Processing Systems
Having Different Processing Speeds".

- 5 -

In addition, in the embodiment under consideration, the memory groups 1 has a plurality of zones 4 for storing information pages before display, these zones being called page memories, and has other zones 5 for storing character information, these being called character generators.

Zones 4 and buffer memory 2 are random access memories while zones 5 can be read only memories.

In addition to these memories for storing the data to be displayed, the memory group includes, according to the invention, zones for memorizing the data to be reproduced in an audible form, and, in particular, in a spoken form. These are zones 6 which constitute the memories for the pages to be spoken.

The number of zones 6 is equal to the number of zones 4 of the memories for the pages of the information to be displayed.

In this embodiment, there are three of these zones.

Memories 4 for the pages to be displayed, and memories 6 for the pages to be spoken, are connected to table 7 for text-word association.

Table 7 includes as many rows 8 as there are text memories 4.

Each row 8 includes register 9 containing the number of the corresponding page, register 10 containing the address of the text of this page to be displayed, and at least one address register corresponding to the language in which the commentaries are to be spoken.

In this embodiment, it is contemplated that the commentaries are to be spoken in any of three languages, namely, French, English, and German.

The table 7 includes, therefore, in each row, three address registers 10a, 10b, 10c for the three languages.

5 The association table 7 is connected to start of page register 11, which is a counter in register stack 12, this latter being a part of control circuit 13.

This control circuit effects the control of the contents of the buffer memory in a manner
10 described in the above mentioned French patent application N° 80 26 393.

As indicated in said application, the start of buffer memory register DMT and of buffer memory register FMT, as well as other registers for storing
15 addresses relating to other zones of memory group 1, make up a part of the register stack 12.

The register stack 12 is connected to a network including data bus 14 connected to a data register 15, which latter is itself connected by bus
20 16 to memory 1, to didon circuit separator 17 for the multiplex signal 3, and to means 18 for processing the typewritten and graphical data to be displayed on screen of tube 19, a decoder 20 being inserted between means 18 and tube 19.

25 The control circuit 13 also includes accumulator 21 connected to address bus 22 which communicates with address register 23 which is itself connected with the address input of the memory group 1 by means of bus 24.

30 State register 25, which also makes up a part of the control circuit 13, is connected to central processing unit 26.

The teletext display device also includes arithmetic and logic unit 27 which processes the

addresses before they are sent to memory group 1,
the addresses, on the one hand, from register stack
12, and accumulator 21 over bus 28 and, on the other
hand, from device 18 over bus 29 which carries the
5   addresses, in ASCII code, of the characters stored
in character generators 5 of memory group 1. A time
base circuit 30 connected to the didon separator
device 17 and to other elements of the display device
by bus 31 provides for the synchronization of the
10   operations.

Control register 37 provides the control
of the data circulating over the buses and direct
access device 33, which is of the transparent type, effects
the control of access to memory group 1.

15   A direct access device of this type is
described in French patent application N° 77 31 140
filed October 17, 1977 by the instant applicant.

The direct access device 33 is associated
with control logic 34.

20   There is also included keyboard 35 which
includes a key 36 for row by row vertical scanning in
typographic mode, a key 37 for line by line vertical
scanning of the image frame in typographic or graphic
mode, and a key 38 for horizontal scanning of the
25   image, and includes, in addition, keys 39 to 41 for
selection of the language in which it is desired
that the commentary accompanying the text be spoken.

Before continuing with the description of
the receiving station of Fig. 1, there is given the
30   following information as to the encoding and decoding
technique for the audible signals.

The manner of reconstitution of the audible
signals is a function of the manner in which these
signals are encoded upon emission.

- 8 -

In the instant example, the word encoding
technique utilizes a dictionary of stored audible
segments and utilizes certain transitions between
classes of audible segments. These elements are either
5   simply concatonated or connected according to very
simple rules.

Each language can be reproduced from a
specific group of elementary sounds.

All of the words of a given language are
10   constructed from these base sounds. The selection of
audible segments and transitions is effected from the
phonological transcription (in the form of phonemes)
of the phrase to be synthesized.

The phonetic rules which operate as a function
15   of the phonetic context, in which is placed the
phoneme under consideration, allows for the selection
of the elementary sounds.

This phase is called "phonetic-phonemic
transcription". The phonemes are themselves obtained
20   from the orthographic text by means of an algorithm
of phonemic-ortographic transcription.

According to the phonetic-phonetic-
orthographic code which is transmitted, these
transcription algorithms are or are not read in via
25   the antiope terminal.

Each audible segment of the dictionary thus
constituted represents a sound adapted to be repro-
duced by a synthesizer circuit to which there is
periodically sent a structure of data or elements.

30   Each audible segment corresponds therefore
to a limited sequence of elements. The coding
technique of the prosody consists in associating,
to each element, two parameters, which parameters
are, the duration of the element and the frequency

of the fundamental sound of this element.

It is noted that the group of elements associated with each audible segment is invariable while the parameters representing the prosody vary with

5 each word to be spoken.

The constitution of a word can be represented as follows :

$$\underline{/}^-\text{WORD}\_\underline{7} = \underline{/}^-(\text{AUDIBLE SEGMENT}_1)\ldots(\text{AUDIBLE SEGMENT}_n)\_\underline{7}$$

10 An audible segment has the following constitution :

$$\text{AUDIBLE SEGMENT} = \Big\{\underline{/}^-(\text{Element}_1)+\text{duration}_1+\text{fundamental}_1\_\underline{7}..$$
$$\ldots\underline{/}^-(\text{Element}_k + \text{duration}_k + \text{fundamental}_k\_\underline{7}\Big\}$$

If one takes into account the fact that the

15 list of elements associated with an audible segment is invariable, it will be noted that each word to be spoken with a desired intonation is, in the end, characterized by the expression :

$$\underline{/}^-\text{number of audible segment} + \sum (di,pi)\_\underline{7}$$

20 in which :

d is the duration of the audible segment elemnt;

p is the fundamental frequency of the element;

i is the order of the element in the audible segment.

25 The transmission of a word becomes thus the transmission of the above sequence, assuming that, upon reception, there will be in the memory the list of the elements associated with each audible segment.

Returning to the architecture of the part

30 of the device which effects the processing of the signals for generating the commentary associated with the displayed pages, there is seen a language memory 42 made up of three memories 42a, 42b, 42c associated respectively with address registers 10a, 10b, 10c for the

35 three languages.

Each of these language memories includes as
many rows 43 as exist sound segments necessary for the
speaking of the language in question.

Each line 43 contains a number 44 of sound

5   segments, as many couples 45 of terms (di, pi) as
there are elements or "frames" in the corresponding
audible segment.

The couples of terms (di, pi) determine the
intonation of the phrases of the commentary.

10   The couples (di,pi) are the attributes of
the elements making up the audible segments which are
stored in memories 46a, 46b, 46c in the form of tables
of elements constituting the audible segments for the
respective languages.

15   Each of these memories, which are, in this
case, read only memories, are adapted to store the
accoustic parameters or coefficients of filters of
formation of the language without intonation which,
in association with the data memorized in the language

20   memory 42a, 42b, or 42c, controls a word synthesizer
47 of the LPC type (linear predictive coding) of the
TMS5220 type. In the present example, the language
memories 42a to 42c and the tables of elements consti-
tuting the audible segments 46 are connected to

25   synthesizer 47 by means of multiplexor 48.

The output of the synthesizer is connected
to a loud speaker 49 which, if the cathode ray tube
is a television tube, can also be the speaker of this
television.

30   Instead of having as many read only memories
46a to 46c as there are programmed languages, one can
also have a single RAM loaded in a dynamic manner with
the audible segments corresponding to the desired
language, and this under the prior control of the

language selection which causes the input of the audible segments into the RAM.

The loading of the RAM is controlled by the central processing unit 26 which "awaits" a transmitted page containing the element making up the audible segments.

Upon the detection of this page, the RAM is loaded with audible segments of the selected language.

The changing of the language is effected by removing from the RAM the elements of the audible segments of the previously utilized language and in replacing them in an analogous manner by other elements.

The functioning of the device which has been described is as follows.

As is seen in Fig. 2, a page of text to be transmitted is subdivided into rows to which are assigned numbers.

As an example, there is assigned the numbers 0 to 24 to the rows of the written text to be transmitted and displayed.

The numbers 30 to 40 are assigned, for example, to the rows in which are written the codes corresponding to the word in French.

The numbers 41 to 50 are assigned to the rows in which are inscribed the codes corresponding to the word in English.

The numbers 51 to 60 are assigned to the rows in which are written the codes corresponding to the word in German.

The information in a page includes :
- A code of the start of the page
- A code of the page number
- A heading.

Thereafter, each row is marked by a code

of the start of the row followed by a code of the
number of the row which is adapted to provide a
recognition of the association of the row to a written
text or to a spoken text.

5   The number of the code of the row is in turn
followed by codes of the information itself correspon-
ding to the text to be displayed or spoken.

      Upon the reception of the signal codes
thus constituted, the central processing unit 26,

10 which is preferably a microprocessor TMS7040, procedes
with the discrimination of the rows of the written text
and the rows of the spoken text and their storing in
memories 4 or 6 according to their nature.

      As this stage of functioning, if it is desired

15 .to display initially the text and to follow this display
by a corresponding commentary, there is emitted the
rows of each page in sequence in their natural order.

      In such a case, the rows 0 to 24 corresponding
to the written text are initially received at the

20 input 3 and immediately displayed. The data relating
to the word are only received thereafter.

      If it is desired to display simultaneously
with the sound, there is intermingled upon transmis-
sion the numbers of the rows of the written text and

25 of sound according to a predetermined sequence.

      Table 7 of text-word association for a page
already decoded, the data of which are stored in written
page memory 4 and spoken memory 6, is continually
updated at the same time that the arranging of the

30 word and text signals in the respective memories is
carried out.

      If the device functions in several languages,
there is a transmission of the group of rows of each
page corresponding to the texts spoken in all of the

35 languages.

From this it results that all of the program-
med languages are available at each instant at the
reception level.

The selection of the language in which the
5   commentary is to be spoken is effected by the user who
activates one of the keys 38 to 41 of the keyboard 35.

The central processing unit 26 controls the
association table 7 so that this can activate the
language memory 42a to 42c and the table of constituting
10   elements 46a to 46c corresponding to the selected
language.

The data corresponding to a page to be spoken
contained in one of the memories 6 are then transmitted
to the corresponding language memory 42a to 42c and
15   the audible segments formed of elements constituted
by parameter couples (d_i, p_i) of intonation are
combined with the corresponding constituting elements
contained in the table 46a to 46c associated with the
memory of the language in question, in multiplexor 48
20   which controls, in its turn, the synthesizor 47.

The sounds thus synthesized are emitted by
the loud speaker in a predetermined sequence with the
text displayed on the screen of the cathode ray tube 19.

In the following list, there is indicared
25   as an example the elements capable of effecting the
functions of the various circuits in Fig. 1b.

Memory 6	
Memory 7	TMS 4164
Memories 42a to 42c	
30	Separator circuit 17	TMS 3534
	SN 96533
Synthesizor 47	TMS 5220
Multiplexor 48	Function of microprocessor
	26 (Fig. 1A) TMS 7040.

35   All of these components are manufactured and
sold by the applicant.

<u>C L A I M S</u>

1. A method for coordinating the audition of spoken messages with the visualization of messages, formed of pages, on a display device, characterized in that it

5   includes receiving the image signals and spoken signals, coded by means of similar codes, the codes corresponding to a written text being recognizeable by a first series of row numbers of the page of the written text assigned to these codes, while the codes corresponding to the spoken

10   message are recognizeable by at least a second series of row numbers of the page assigned to these codes of the spoken message, carrying out a discrimination between the codes of the written text and the codes of the spoken message, storing the codes of the written text in first

15   memory grouping means, storing the codes of the spoken message in second memory means, associating the codes of the written text with the codes of the corresponding spoken message by means of an association table, reconstituting the spoken message from the codes contained in the second

20   memory means by causing the audible emission thereof according to a predetermined sequence, with the display of the corresponding written text on the screen of a cathode ray tube.

2. A method according to Claim 1 for the audi-

25   tion of a spoken message in one of several programmed languages, characterized in that the codes corresponding to the same message to be given in each language are recognizeable by a series of row numbers of the page assigned to the programmed languages respectively.

30   3. A Method according to Claim 2 characterized in receiving the codes corresponding to a spoken message in all of the programmed languages, storing the codes corresponding to the spoken messages in each language in a corresponding memory of said second memory means, and

35   in effecting the reconstitution of the spoken message in

one of the selected languages from the codes contained in
said corresponding memory of the second memory means.

4. A method according to any one of the preced-
ing claims, characterized in receiving the codes of the
spoken message corresponding to the intonation elements
of the sound segments forming the message, and combining
the intonation elements with the elements constituting
said sound segments, stored in said second memory means.

5. A device for coordinating the audition of
spoken messages with the visualization of messages, made
up of pages, on a display device, characterized in that
it includes means (4, 5) for memorizing the codes corres-
ponding to the written text to be displayed, means (6)
for memorizing the codes corresponding to the spoken
message to be reproduced, means (7) for associating the
text and the speech corresponding to the codes stored in
said memorizing means (6), means (42a - 42c) for memori-
zing the intonation elements (di, pi) of the sound
segments forming the received message, means (46a - 46c)
for memorizing the elements constituting the language
in which the message is received, means (48) for combining
said intonation elements with said elements constituting
the language, and means (47) for synthesizing the message
to be reproduced under the control of said combining means.

6. A device according to Claim 5, characterized
in that said means (7) for associating the text and the
speech include a table having a number of lines (8) equal
to the number of page memories (4) of the means for memo-
rizing the text codes, each of the lines (8) including a
first register (9) containing the corresponding page num-
ber (4), a second register (10) containing the address of
the text of this page to be displayed, at least a third
address register (10a, 10b, 10c), corresponding to a
language in which the spoken message is to be given.

7. A device according to Claim 6, characterized

in that said third address register is made of as many
address registers (10a, 10b, 10c) as there are programmed
languages in which the spoken message could be given.

5   8. A device according to one of the Claims 5 to
7, characterized in that said means for memorizing the
intonation elements (di, pi) of the sound segments include
at least one language memory (42a - 42c) including as many
lines (43) as there are sound segments necessary for the
pronunciation of the language considered, each line (43)
10   containing a sound segment number (44) and as many pairs
of terms (di, pi) of duration and fundamental frequency
as there are elements in the corresponding sound segment,
said pairs of terms (di, pi) determining the intonation
in which the sound segment is to be made audible.
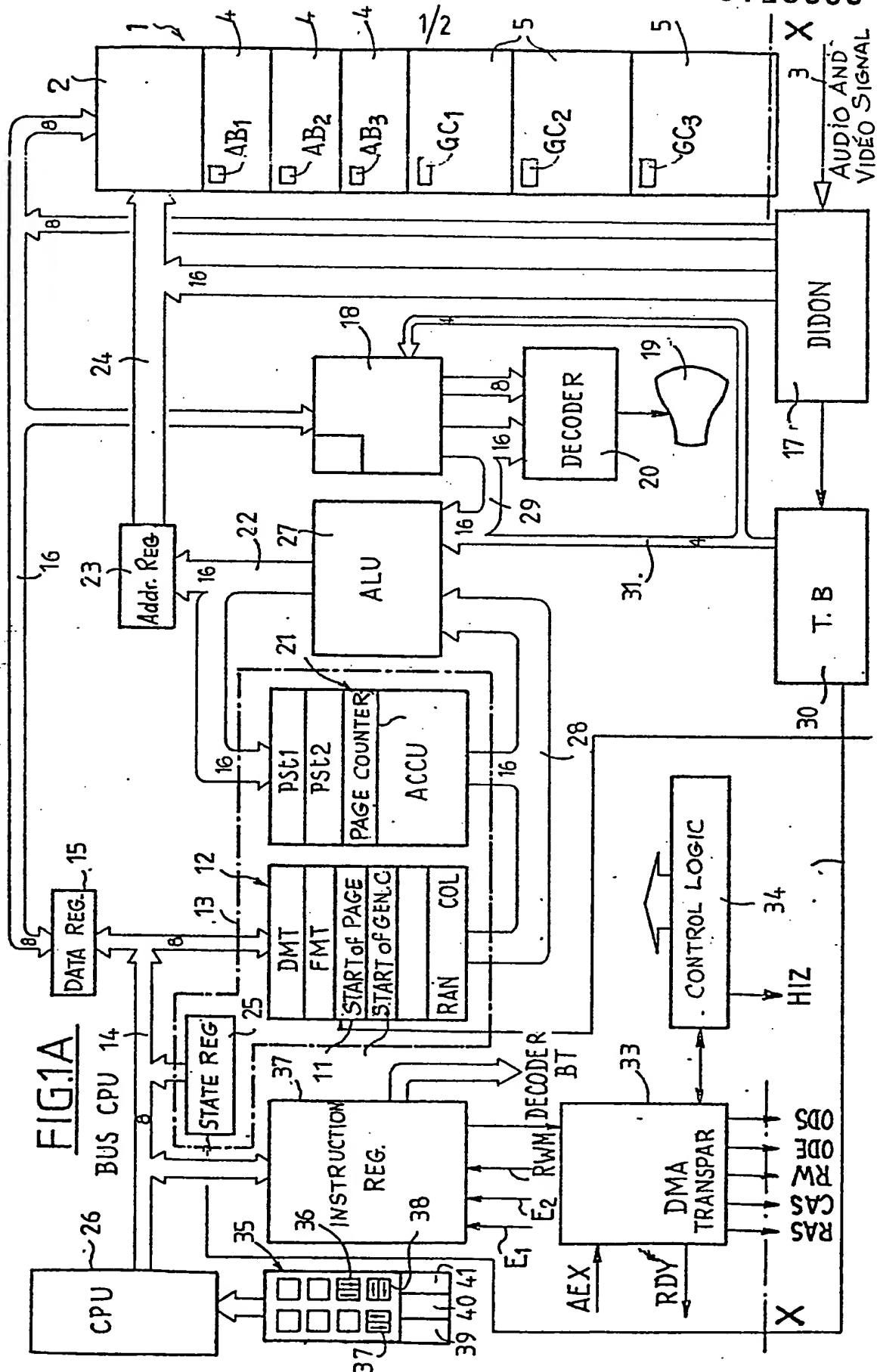
15   9. A device according to Claim 8, characterized
in that said means for memorizing said intonation elements
include as many language memories (42a - 42c) as there are
languages programmed in the device.

10. A device according to any one of the Claims
20   5 to 9, characterized in that the means for memorizing the
elements constituting the language in which the spoken
message is to be made audible, includes as many memories
containing the parameters of said elements constituting
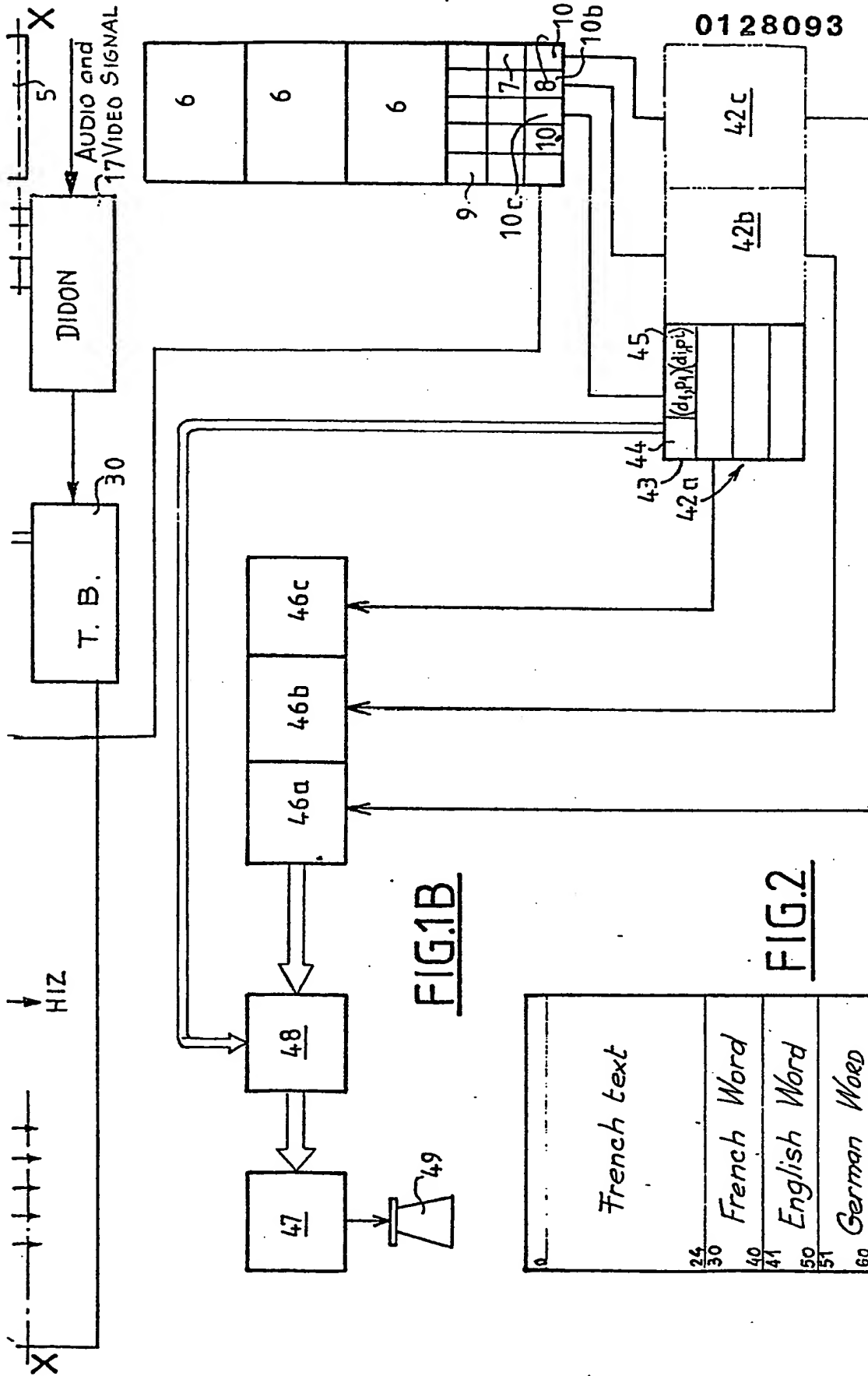the language as there are programmed languages.

25   11. A device according to Claim 10, characteriz-
ed in that said memories are read only memories.

12. A device according to any one of the Claims
5 to 9, characterized in that the means for memorizing the
elements of the language in which the message is to be
30   made audible includes a random access memory (RAM) adapted
to store, in a dynamic manner, under control of a language
selection command (39, 40, 41), the parameters of said
elements constituting the selected language.

0128093

FIG1A

0128093

FIG.1B

FIG.2

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication. where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl. ³)
Y	DE-A-3 021 647 (SIEMENS AG) * page 6, line 22 - page 12, line 4 *	1	H 04 N 7/08
Y	EP-A-0 071 716 (TEXAS INSTRUMENTS INCORP.) * page 5, line 15 - page 6, line 21; page 11, line 1 - page 13, line 12 *	1	
A		4,5	
A,D	EP-A-0 055 167 (TEXAS INSTRUMENTS FRANCE)		

TECHNICAL FIELDS SEARCHED (Int. Cl. ³)

H 04 N

The present search report has been drawn up for all claims

Place of search THE HAGUE	Date of completion of the search 10-09-1984	Examiner VERSCHELDEN J.